



Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Fair bandwidth sharing and delay differentiation: Joint packet scheduling with buffer management

Xiaobo Zhou^{a,*}, Dennis Ippoliti^a, Liqiang Zhang^b

^a Department of Computer Science, University of Colorado at Colorado Springs, CO 80918, USA

^b Department of Computer and Information Sciences, Indiana University South Bend, IN 46615, USA

ARTICLE INFO

Article history:

Received 7 March 2008

Received in revised form 5 August 2008

Accepted 13 August 2008

Available online 22 August 2008

Keywords:

Proportional differentiation

Fair bandwidth sharing

Multi-dimensional Quality of service

Packet scheduling

Buffer management

Feedback control

ABSTRACT

Packet delay and bandwidth are two important metrics for measuring quality of service (QoS) of Internet services. Traditionally, packet delay differentiation and fair bandwidth sharing are studied separately. In this paper, we first propose a generalized model for providing fair bandwidth sharing with delay differentiation, namely FBS-DD, at the same time. It essentially aims to provide multi-dimensional proportional differentiation with respect to both QoS metrics. We design size-based packet scheduling schemes that take both packet delay and packet size into scheduling considerations, without assuming admission control or policing. Furthermore, we propose a PID control-theoretic buffer management scheme. The packet scheduling with buffer management approach provides delay and bandwidth differentiation in an integrated way, while existing approaches consider delay and loss rate differentiation as orthogonal issues. It enhances the flexibility of network resource management and multi-dimensional QoS provisioning. It is capable of self-adapting to varying workloads from different classes, which automatically builds a firewall around aggressive clients and hence protects network resources from saturation. Extensive simulation results by the use of trace files demonstrate that the packet scheduling schemes can provide predictable fair bandwidth sharing with delay differentiation at various situations. The control-theoretic buffer management scheme further improves the controllability.

© 2008 Published by Elsevier B.V.

1. Introduction

Differentiated Services (DiffServ) is one of the major efforts to meet the demand of provisioning different levels of quality of service (QoS) on the Internet so as to support different types of network applications and various user requirements. It aims to provide differentiated services between classes of aggregated traffic flows within a router, rather than offer QoS guarantees to individual flows [1]. To receive different levels of QoS, packets are assigned with different service types or traffic classes at the network edges. DiffServ-compatible routers in the network core perform stateless prioritized packet forwarding or dropping, called “per-hop behaviors” (PHBs), to the classified packets. Due to its per-class stateless processing, the DiffServ architecture exhibits good scalability. Its provisioning is an active research topic [5,7,9,10,17,21–25].

There are two basic schemes to DiffServ provisioning. Absolute DiffServ aims to provide statistical assurances for a class's received performance measures, such as a minimum service rate or maximum delay. Relative DiffServ is to quantify the quality spacings between different classes. The proportional differentiation model,

proposed by Dovrolis, *et al.* [4], is a popular relative DiffServ model. It aims to provide per-class QoS level in proportion to the pre-specified differentiation parameters of the classes, independent of those class workloads. Delay and bandwidth are two important QoS metrics considered in the model. The algorithms for proportional delay differentiation (PDD) consider lossless and work-conserving packet scheduling [4–6,11,13,14,17,18]. When the overall workload of classes is close to or exceeds the link bandwidth capacity, the algorithms for proportional bandwidth differentiation aim to enforce that the ratio of the loss rates of two classes be proportional to the ratios of their differentiation parameters [3,7,21,23]. However, most of those algorithms consider delay differentiation and bandwidth differentiation as orthogonal issues.

While the PDD model is excellent due to its delay proportional-fairness to clients, it is insufficient and might be unfair from the perspective of the network resource providers. It is because the model does not consider another important issue, fair bandwidth sharing. Fair bandwidth sharing is a classic issue. Its short-term behaviors were originally studied as fair queueing [2]. While those PDD algorithms can ensure that experienced delay of different classes be proportional, there is no assumption nor guarantee on the fair bandwidth sharing, be in short term or in long term. Consider two traffic classes (Class-1 and Class-2) with the pre-specified differentiation parameters 2 and 1, respectively. Consider the

* Corresponding author. Tel.: +1 719 262 3493.

E-mail addresses: zbo@cs.uccs.edu (X. Zhou), liqzhang@iusb.edu (L. Zhang).

scenario that Class-1's workload is 80% of the link capacity and Class-2's workload is 5% of the link capacity. According to the proportional delay differentiation model, the ratio of the average packet delay of Class-1 to that of Class-2 would be 1 to 2. However, the workload of Class-1 is 16 times of that of Class-2 while their differentiation parameter ratio is only 2 to 1. The scenario illustrates that the current workload-independent proportional differentiation model can be very unfair to some network traffic. Even worse, some aggressive or malicious clients can utilize this unfairness and weak controllability to attack the network resources.

Note that we do not intend to deny the merit of the PDD model. Essentially, it considers the single-dimensional QoS provisioning with respect to delay. It needs the support of admission control schemes that shape the traffic according to the service level agreements or some adaptive schemes that promote the differentiation parameters dynamically according to the workload conditions. Generally, the pre-specified differentiation parameters are used by the network operators to control the quality spacings between the multiple classes. They are often associated to the differentiated pricing, say proportionally. But the model is insufficient when multiple QoS metrics exist and multi-dimensional QoS provisioning should be considered.

Given that both bandwidth and delay are important metrics for measuring QoS of Internet services, we need to consider fair bandwidth sharing and delay differentiation at the same time. The primary contributions of our work are:

1. We propose a generalized model, FBS-DD, for providing fair bandwidth sharing with delay differentiation at the same time. It is to ensure that the ratio of the average delay of two classes normalized by their achieved bandwidth ratio be proportional to the pre-specified differentiation parameters. It essentially aims to provide multi-dimensional proportional differentiation with respect to both QoS metrics, packet delay and bandwidth. One uniqueness is that the delay differentiation and loss rate differentiation are integrated with traffic policing capabilities for providing better controllability to network operators and fairness to clients.
2. We design size-based packet scheduling algorithms for FBS-DD provisioning, modified from the waiting-time priority (WTP) algorithms which are excellent schedulers for performing proportional delay differentiation. Two VPS (various packet size) algorithms take both packet size and packet delay into consideration in packet scheduling. For packets with the uniform size, the VPS schemes are reduced to UPS (uniform packet size) schemes.
3. We further study the performance controllability with control-based buffer management. When the overall workload of the classes is below the link capacity, the FBS-DD model actually is to achieve the proportional delay differentiation weighted by the workloads of the classes in the long term. When the overall workload of classes is beyond the link capacity so that there will be packet loss, the FBS-DD model is to achieve the proportional delay differentiation weighted by the experienced bandwidth ratio of the classes. This is however a non-trivial issue. We propose a PID control-theoretic buffer management scheme to further provide proportional loss rate differentiation along with the FBS-DD provisioning. The controller enhances the controllability of network resource management.
4. We conduct extensive performance evaluation based on the simulation by the use of Bell Labs-I IP trace files. Results show that the proposed scheduling and buffer management schemes are capable of self-adapting to varying workloads of different classes. They automatically build a firewall around aggressive clients and protect network resources from saturation.

Our work is to address the integration of traffic policing with proportional differentiation. The study provides insights to the multi-dimensional differentiated services provisioning. The structure of the paper is as follows. In Section 2, we review existing packet scheduling and dropping algorithms for proportional differentiation provisioning. Section 3 presents the FBS-DD model with packet scheduling and buffer management schemes. Section 4 focuses on the performance evaluation. Section 5 concludes the paper.

2. Related work

Fair bandwidth sharing was initially studied as fair queueing [2], which aims to allow each flow passing through a network device to have a fair share of network resources. There are classic mechanisms for achieving the short-term per-flow fairing sharing, see PGPS [20] and a random scheme in [16] for examples. There is also recent study on fair load sharing in multipath communication networks [12]. In the context of DiffServ, the QoS provisioning is concerned with per-class behaviors. The FBS-DD model considers the long-term fair bandwidth sharing with delay differentiation.

2.1. Packet scheduling for proportional delay differentiation

Delay differentiation in packet networks is an active research topic. The PDD model is to provide differentiated delay services among traffic classes [4,5]. A class is assigned a delay differentiation parameter. The packet scheduler of a router aims to keep the ratio of average delay of a higher priority class to that of a lower priority class equal to the pre-specified value. The existing PDD algorithms can be classified into three categories [25].

Rate-based packet scheduling algorithms adjust service rate allocations of classes dynamically to meet the proportional delay differentiation constraints [4,13,14]. BPR [4] adjusts the service rate of a class according to its backlogged queue length so that the class service rates are proportional to the corresponding ratios of class loads. JoBS [14] allocates the service rate of a class based on delay predictions of its backlogged traffic. It forms the service rate allocation into an optimization problem when the system is heavily loaded. The objective of JoBS is to enforce absolute delay and loss constraints. The accuracy of the rate-based algorithms over the delay ratio is unfortunately dependent of class load conditions [4]. This is because they rely on the relationship between queueing delay and service rate for a backlogged queue. However, the class load distribution on a router tends to change quickly. This limits the applicability of the algorithms.

Time-dependent priority packet scheduling algorithms adjust the priority of a backlogged class according to the experienced delay of its head-of-line packet. In WTP [5], the priority of a backlogged class is set equal to the waiting time of the head packet normalized by its differentiation parameter dynamically on departure of each packet. A packet of a backlogged class with the highest priority will be forwarded next. Albeit simple, WTP implements the PDD model only when the system utilization approaches 100%. In AWTP [11,6], the control parameter of class is adjusted according to its class load dynamically. Moreover, a necessary condition was derived in [11], with respect to the class load conditions, for feasible WTP control parameters to achieve desired class delay ratios. It has better accuracy and adaptivity, in comparison with WTP, in both short and long timescales.

Little's law-based packet scheduling algorithms correlate the average queue length to the average arrival rate and the average queueing delay of packets. They control the actual delay ratio between two different classes by equalizing their normalized queue lengths with the pre-specified delay differentiation parameters.

Table 1
A summary of the representative PDD packet scheduling algorithms

Algorithms	Packets considered
BPR	backlogged packets
JoBS	backlogged packets
WTP	head-of-line packet
AWTP	arrived and head-of-line packets
PAD	departed packets
HPD	departed and head-of-line packets
MDP	departed and backlogged packets
LAD	departed and backlogged packets

The equalization process is a feedback control process. They differ in the way of average delay calculation. PAD [5] considers the average delay of departed packets in the time window only. In long timescales, it is capable of achieving the PDD model. It, however, exhibits a pathological behavior in short timescales. HPD [5] takes into account both the average delay of departed packets and the delay of the head-of-line packet simultaneously. MDP [18] considers the delays of all arrived packets of each class in a time window and the estimated delays of backlogged packets in future. It achieves the class delay ratio with small variations. LAD [22] monitors the arrival rate and the cumulative delays of packets of each class. It selects the next packet according to their transient queueing properties so as to achieve the desired delay ratios in both short and long timescales. Table 1 summarizes those representative algorithms.

2.2. Packet dropping for loss rate differentiation

There are a number of interesting differentiated buffer management and packet dropping schemes for loss rate differentiation. PLR droppers in [3] aim to provide proportional loss rates to different traffic classes according to their differentiation weights. JoBS in [14] extends the proportional loss rate model by providing both absolute loss and delay guarantees and proportional differentiations. The proportional differentiation constraint is relaxed to satisfy the absolute constraints when the two sets of constraints cannot be simultaneously satisfied. BRD dropper in [7] seeks to minimize the loss rate differences between classes subject to the absolute loss constraints and the relative loss constraints. HPPD in [23] aims to reduce the retransmission cost of the dropped packets for congestion mitigation by hop-count based differentiated packet dropping.

Those dropping schemes are able to achieve their differentiation objectives. But the schemes, exception of JoBS [14], consider delay differentiation and bandwidth differentiation as orthogonal issues. There are two significant differences between JoBS and our work. JoBS's goal is to support both absolute and relative Diff-Serv. It executes an optimization on every packet arrival. Ours is to provide multi-dimensional QoS with respect to both bandwidth sharing and delay differentiation. Packet scheduling is lightweight. Second, JoBS is rate based, which assumes a fluid-flow interpretation of traffic. It hence needs the support of scheduling algorithms that closely approximate the fluid-flow schedulers with rate guarantees. Our work follows a practical per-packet scheduling discipline. Furthermore, it adopts a unique control-theoretic buffer management scheme for the controllability improvement.

3. The FBS-DD Model and algorithms

We consider the work-conserving packet scheduler that serves N queues, one for each class. The FBS-DD model is to maintain the multi-dimensional QoS spacing of two classes with respect to their delay ratio (D_i/D_j) normalized by their bandwidth sharing ratio

(B_i/B_j) be proportional to their pre-specified differentiation parameters δ_i and δ_j . That is,

$$\frac{D_i(T, T+t)}{D_j(T, T+t)} \frac{B_j(T, T+t)}{B_i(T, T+t)} = \frac{\delta_i}{\delta_j}, 1 \leq i, j \leq N \quad (1)$$

for time intervals $(T, T+t)$ where t is the monitoring timescale.

Note that the lower average delay or higher bandwidth sharing represents higher QoS. FBS-DD is essentially a fair tradeoff between PDD provisioning and FBS provisioning. When two classes experience the same bandwidth, the model is reduced to PDD model. When two classes experience the same average delay, the model is reduced to FBS model.

3.1. Packet scheduling schemes

3.1.1. VPS-TWP: Throughput normalized waiting time priority scheduling

First, we consider the general case, that is, packets from a class have various sizes and different classes may have different packet size distributions. We revisit the time dependent priority scheduling discipline and design the VPS-TWP scheme, which focuses on the instantaneous behavior. The time dependent priority scheduling was first studied in queueing foundations. It was later studied by Dovrolis *et al.* in WTP [5] for PDD provisioning. We describe VPS-TWP as the throughput normalized waiting time priority scheduling algorithm for FBS-DD provisioning.

At the beginning of scheduling, $TWP_i = \infty$ for $1 \leq i \leq N$. Suppose that class i is backlogged at time t , $s_i(t)$ is the size of the packet at the head of the class i at t , and that $w_i(t)$ is the head waiting time of class i at t , i.e., the waiting time of the packet at the head of the class i at t . We define the throughput normalized head waiting time of class i at t as

$$TWP_i(t) = \frac{w_i(t)}{\delta_i s_i(t)}. \quad (2)$$

Every time a packet is to be transmitted, the VPS-TWP scheduler selects the backlogged class j with the maximum throughput normalized head waiting time,

$$j = \arg \max_{i \in G(t)} TWP_i(t), \quad (3)$$

where $G(t)$ is the set of backlogged classes at time t . Tie breaks by the use of priority. The throughput of class j is increased by the size of the transmitted packet. Its throughput normalized head waiting time will be minimized as its packet delay will not increase any more. VPS-TWP attempts to minimize the differences between the bandwidth normalized waiting times of successively departing packets. It essentially aims to achieve instantaneous FBS-DD.

Next, we consider a special case, that is, all packets have the uniform size. The experienced bandwidth ratio of classes i and j is given as $B_i(T, T+t)/B_j(T, T+t) = b_i/b_j$ where b_i and b_j are the number of packets departed in the interval $(T, T+t)$, essentially the throughput of the classes in the interval. When all packets have the uniform packet size, VPS-TWP is reduced to UPS-TWP. The throughput normalized head waiting time of class i at t is calculated as

$$TWP_i(t) = \frac{w_i(t)}{b_i \delta_i}. \quad (4)$$

3.1.2. VPS-TAD: Throughput normalized average delay scheduling

While VPS-TWP focuses at the instantaneous behavior, we design the VPS-TAD scheme which focuses on the long-term behavior. (1) can be rewritten as

$$\frac{D_i(T, T+t)}{\delta_i B_i(T, T+t)} = \frac{D_j(T, T+t)}{\delta_j B_j(T, T+t)}. \quad (5)$$

It is interpreted that the throughput normalized average delay (TAD) must be equal in all classes. That is $TAD_i = TAD_j$. Note that given a same interval, bandwidth sharing ratio of two classes is the same as the throughput ratio. The VPS-TAD scheme, tailored from PAD [5], aims to equalize the throughput normalized average delays among all classes so as to achieve the FBS-DD goal.

Let $G(t)$ is the set of backlogged classes at time t , $L_i(t)$ be the sequence of class i packets that were transmitted during the interval $(T, T+t)$, d_i^m be the delay of the m th packet in $L_i(t)$, and s_i^m be the size of the m th packet in $L_i(t)$. Assuming that there was at least one packet transmitted from class i during interval $(T, T+t)$, the throughput normalized average delay of class i at t is

$$TAD_i(t) = \frac{D_i(T, T+t)}{\delta_i B_i(T, T+t)} = \frac{1}{\delta_i \sum_{m=1}^{|L_i(t)|} s_i^m} \sum_{m=1}^{|L_i(t)|} d_i^m, \quad (6)$$

where $|L_i(t)|$ is the number of packets in $L_i(t)$.

At the beginning of scheduling, $TAD_i = \infty$ for $1 \leq i \leq N$. Suppose that a packet is to be transmitted at time t . VPS-TAD selects the backlogged class j with the maximum bandwidth normalized average delay,

$$j = \arg \max_{i \in G(t)} TAD_i(t). \quad (7)$$

Tie is broken by the priority. The rationale of VPS-TAD is that each time a packet from class j is transmitted, its throughput normalized average delay decreases. This is because its throughput increases by the size of the transmitted packet. The delay of that transmitted packet will not increase any more, and thus the increase to the average packet delay will be minimized. Note that as others work in [5,7,14,22], only the delay of transmitted packets is considered in the model. VPS-TAD therefore attempts to minimize the differences between the throughput normalized average delay of classes. It essentially aims to achieve FBS-DD in the long term. It, however, needs to maintain the state information about the current throughput and average delay per each class.

When all packets have the uniform size, VPS-TAD is reduced to UPS-TAD. The throughput normalized average delay of class i is calculated as

$$TAD_i(t) = \frac{D_i(T, T+t)}{\delta_i B_i(T, T+t)} = \frac{1}{b_i \delta_i} \frac{\sum_{m=1}^{|L_i(t)|} d_i^m}{|L_i(t)|}. \quad (8)$$

3.2. PID Control-theoretic buffer management

When the overall workload is greater than the link capacity, packet loss is inevitable and loss rate becomes the dominant QoS metric. One goal of DiffServ is to enforce the proportional loss rate differentiation between classes. The proposed packet scheduling schemes for FBS-DD provisioning, however, have no control over the loss rate differentiation between classes. We propose a control-theoretic buffer management scheme, to be integrated with the packet scheduling schemes, for the FBS-DD provisioning and proportional loss rate differentiation at the same time. One nice feature of the buffer management based approach is that the packets will be dropped from the tail due to the buffer overflow. This avoids the packet push-out issue and facilitates the packet ordering.

The buffer management is to dynamically allocate the buffer space into a number of virtual mini-buffers, one mini-buffer for one class. The size of a mini-buffer directly affects a class's loss rate. Feedback control theory has been applied to adjust the resource allocation for service differentiation provisioning [10,15]. We propose to use a proportional integral derivative (PID) control-

ler to adjust the buffer allocation. Let l_i be the loss rate of class i . The goal is to ensure that the observed relative loss rate l_i be proportional to the pre-specified QoS parameter δ_i , that is, $l_i/l_j = \delta_i/\delta_j$.

Let L_i be the relative loss rate ratio of class i , that is, $L_i = \frac{l_i}{l_1+l_2+\dots+l_n}$. Let L_i^d be the desired relative loss rate ratio of class i , that is, $L_i^d = \frac{\delta_i}{\delta_1+\delta_2+\dots+\delta_n}$. During the k th sampling period, the relative error is calculated as difference between the desired value and the observed value, that is,

$$e_i(k) = L_i^d(k) - L_i(k). \quad (9)$$

One property of the model is the sum of the relative errors is always zero since $\sum_{i=1}^n e_i(k) = \sum_{i=1}^n (L_i^d(k) - L_i(k)) = 0$. This important property makes it feasible for us to adaptively adjust the buffer allocation for a class independent of the adjustments of other classes while maintaining a constant overall buffer size.

The buffer size allocated to a class is adjusted in proportion to the error between the desired relative loss rate ratio and the observed one. Specifically, the operation of the PID controller is described as follows:

$$s_i(k+1) = s_i(0) + G_P e_i(k) + G_I \sum_{j=0}^{k-1} e_i(j) + G_D \Delta e_i(k). \quad (10)$$

$s_i(k+1)$ denotes the buffer size allocated to class i in the new sampling period. $s_i(0)$ denotes the initial buffer size allocated. The three terms added to $s_i(0)$ denote proportional, integral, and derivative components, respectively. Setting a large proportional feedback gain (G_P) typically leads to faster response at the cost of increasing system instability. The integral controller (G_I) can eliminate the steady-state error and avoid over-reactions to measurement noises. The derivative control (G_D) considers the change of errors in adjusting the buffer size allocation and hence responds fast to errors. The derivative error with class i is calculated as

$$\Delta e_i(k) = e_i(k) - e_i(k-1). \quad (11)$$

4. Performance evaluation

We developed a simulator based on the Click Modular router [8] to study the performance of the packet scheduling schemes and the PID controller based buffer management scheme. For the packet size distribution of different traffic classes, we used Bell Labs-I trace files adopted from the National Laboratory for Applied Network Research [19]. Fig. 1 illustrates the characteristics and the packet size distributions for two Bell Labs-I IP traces. In the following, we first focus on the performance evaluation of the two-class fair bandwidth sharing and delay differentiation. Without loss of generality, let Class-1 be the high priority class and Class-2 be the low priority class. Note that the number of classes for DiffServ often varies from 2 to 3 [5,7,9,25]. Some representative three-class experimental results are presented as the supplements. The first part of the following experiments is to study the impact of the packet scheduling schemes on FBS-DD provisioning when the overall workload is within the link capacity. We considered a lossless model. Computer networks usually over-provision their bandwidth to handle the heavy load and thus the network bandwidth load is normally low in most of time. However, those networks sometimes still face unexpectedly high workloads during unforeseeable events such as terror attacks. Handling sudden spikes or flash crowds is an important problem for Internet services. Therefore, as many others in differentiated service community, we study the capability of FBS-DD model under high workload situations. The second part is on the performance evaluation when the overall workload is beyond the link capacity.

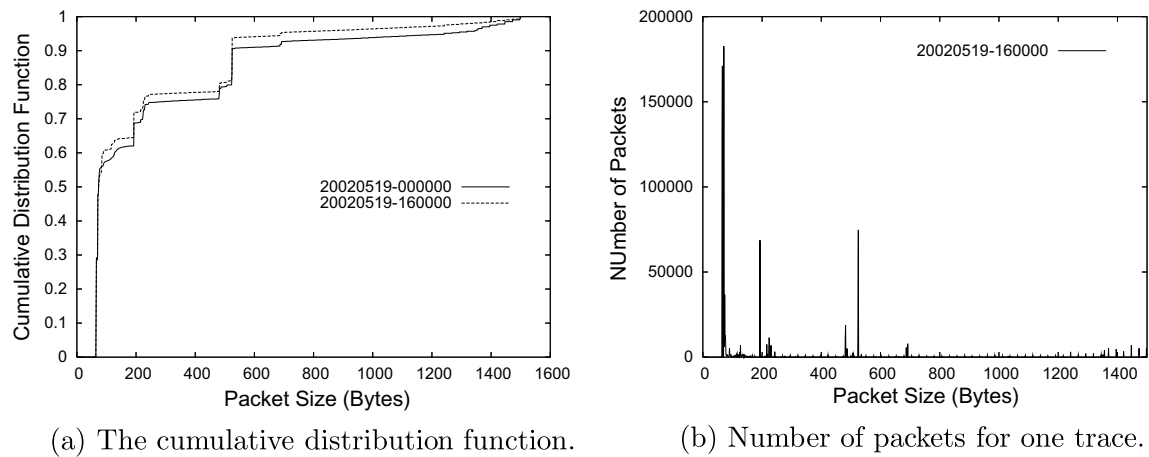


Fig. 1. Packet size distributions of two Bell Labs-I IP traces.

4.1. Fair bandwidth sharing and delay differentiation by packet scheduling

We first study the fair bandwidth sharing and delay differentiation due to the VPS algorithms when the differentiation ratio is fixed and overall workload changes. Figs. 2 and 3 show the performance of the VPS scheduling algorithms when the overall workload changes from 55% to 100%. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) is set to 1:2 and their workload ratio is set to 1:3. Fig. 2 shows the achieved FBS-DD ratio with its 95th and 5th percentiles. Fig. 3(a) shows the achieved delay ratio with its 95th and 5th percentiles. Note that the delay ratio is the FBS-DD ratio normalized by the inverse of the bandwidth ratio. Fig. 3(b) shows the experienced bandwidth ratio of Class-2 to Class-1 with its 95th and 5th percentiles at different time intervals. The results show that the scheduling schemes can achieve the goal of providing fair bandwidth sharing with delay differentiation when the overall workload is greater than 60%. But the variance, as demonstrated by the 95th and 5th percentiles, is a nontrivial issue. It is due to the variance of the packet size distributions and the inter-arrivals. When the workload is low, there is a feasibility issue with the packet scheduling for delay differentiation [5].

We next study the scenarios when the overall workload is fixed but the workload ratio of two classes changes dynamically. We set the target FBS-DD ratio of $\delta_1 : \delta_2$ to 1:3. We fix the overall workload to 80%, and, vary the Class-1's workload from 10% to 90% of the overall workload. Fig. 4 shows the achieved FBS-DD ratio with its

95th and 5th percentiles. It shows the FBS-DD ratio can be achieved as expected. But the variance is high when Class-1's workload deviates from the median value 50%. This is due to the fact that there are too few or too many packets from Class-1, limiting the capability of the packet scheduling schemes.

Fig. 5(a) shows the achieved delay ratio with its 95th and 5th percentiles. We can see that the proposed VPS scheduling schemes can achieve the fair bandwidth sharing with delay differentiation when the workload percentage of the classes changes dynamically. When the Class-1 contributes 75% of the overall workload and the Class-2 contributes 25% of the overall workload, the delay ratio of two classes becomes 1. But on the other hand, the Class-1 achieves 3 times of bandwidth than the Class-2, as illustrated by Fig. 5(b). When the Class-1 contributes 90% of the overall workload (thus, the workload ratio of Class-1 to Class-2 is 9:1), it experiences 3 times higher average delay than the Class-2 does. The results demonstrate the benefit of the FBS-DD model that can make adaptive tradeoff between fair bandwidth sharing and delay differentiation. Note that when the Class-1's workload is only 10% of the overall workload, the expected delay ratio becomes 1:27. This makes it infeasible for the packet scheduling to achieve the target delay ratio.

While both VPS-TAD and VPS-TWP schemes can achieve FBS-DD provisioning from the long-term perspective, they have different behaviors. We studied the behaviors of VPS algorithms in different sampling intervals. Fig. 6 shows the FBS-DD ratios achieved by the scheduling schemes in different sampling intervals. Interestingly, in short sampling intervals, VPS-TAD does not perform well for FBS-DD provisioning. Fig. 6(a) shows that its performance improves as the sampling interval increases. This is explained by the fact that VPS-TAD takes into account the average of a number of packets in the interval. It aims to minimize the differences between the normalized average class delays and thus its performance improves as the sampling interval increases. On the other hand, Fig. 6(b) shows that VPS-TWP achieves desirable FBS-DD ratios when the sampling interval is short and the performance deteriorates as the interval increases. This is due to the fact that VPS-TWP attempts to minimize the differences between the normalized head waiting times. Essentially, it aims to achieve the instantaneous FBS-DD provisioning.

The previous experiments had the FBS-DD ratio fixed. Next we study the performance of VPS algorithms when the FBS-DD ratio changes. The overall workload is fixed to 80% of the link capacity and the workload ratio of Class-1 to Class-2 is fixed to 3:1. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) varies as 1:1, 1:2, 1:3, and so on to 1:16. The results are illustrated in Fig. 7.

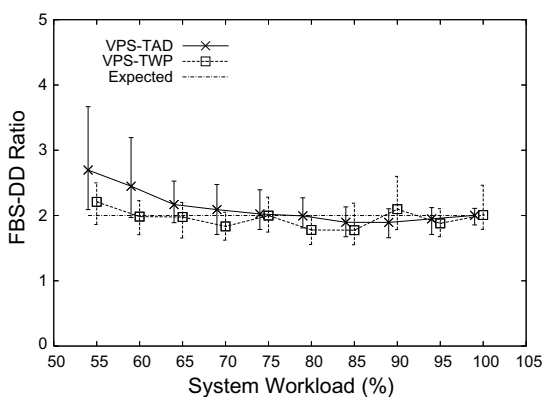


Fig. 2. The FBS-DD ratio due to VPS packet scheduling ($\delta_1 : \delta_2 = 1 : 2$).

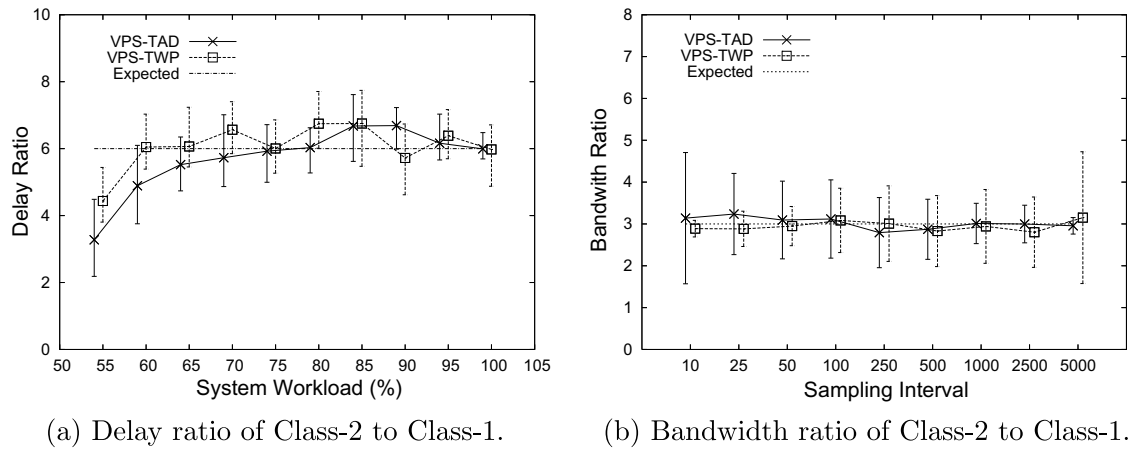


Fig. 3. The delay differentiation ratio and bandwidth sharing ratio due to packet scheduling.

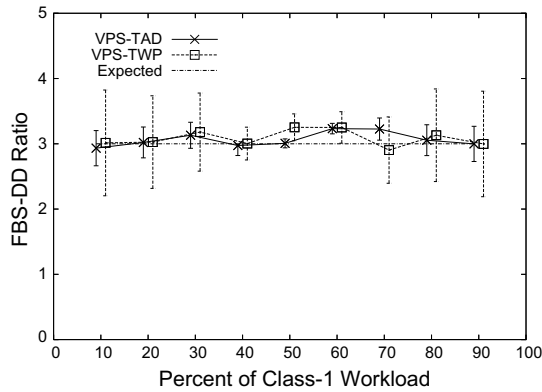


Fig. 4. The FBS-DD ratio due to VPS packet scheduling ($\delta_1 : \delta_2 = 1 : 3$).

We are also interested in the study of the FBS-DD provisioning when one class is obeying its workload promise while the other class changes the workload dynamically. Fig. 8 illustrates the experimental results. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) is fixed to 1:2. The Class-1's workload is fixed to the 60% of the link capacity. The Class-2's workload changes from 20% of the link capacity down to 10% of the link capacity. As the Class-2's workload decreases, both classes' delay decrease as the overall workload decreases. Although the relative delay ratio of Class-1 to Class-2 increases from 3:2 to 2:1 as the Class-2's work-

load is reduced by half, Class-1 is benefitted in terms of the absolute delay. Class-2's delay decreases in a faster pace than that of Class-1. This makes sense since it is reducing its workload. The results demonstrate that the PDD model might not be sufficient in dynamic workload situations. It shows the significant advantage of the FBS-DD model, which considers both bandwidth and delay into multi-dimensional QoS provisioning.

We also conducted experiments for the uniform packet size situations. We found that the UPS schemes can achieve FBS-DD ratios more accurately with smaller variances than the VPS schemes in different situations. It is explained by the fact that there is no packet size variance. Results are omitted as the uniform packet size is very rare.

4.2. Performance of the PID control based buffer management

Previous experimental results have shown that the packet scheduling schemes can simultaneously achieve the fair bandwidth sharing and delay differentiation. But when the overall workload is beyond the link capacity, there will be packet loss and the packet scheduling schemes have no control over the loss rate differentiation between classes. We adopted the VPS-TAD packet scheduling algorithm and conducted following experiments to study the impact of the PID control based buffer management on the loss rate differentiation.

Fig. 9 depicts the impact of the PID control-theoretic buffer management on the controllability of FBS-DD provisioning and

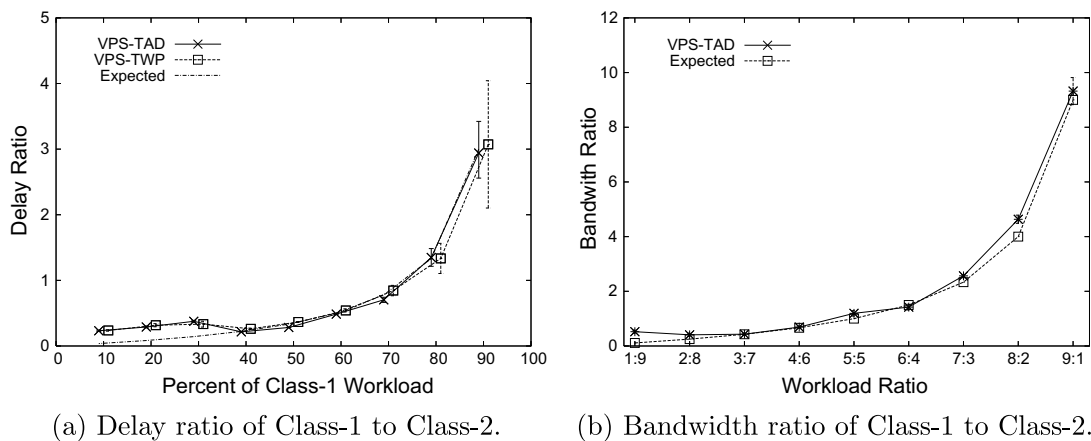


Fig. 5. The performance of VPS scheduling schemes when Class-1's workload changes.

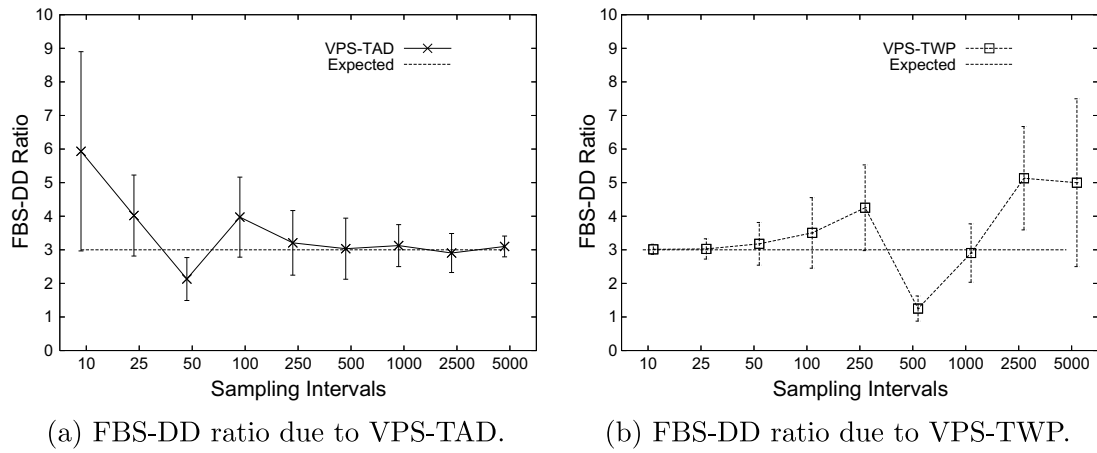


Fig. 6. The behaviors of VPS scheduling schemes in different sampling intervals.

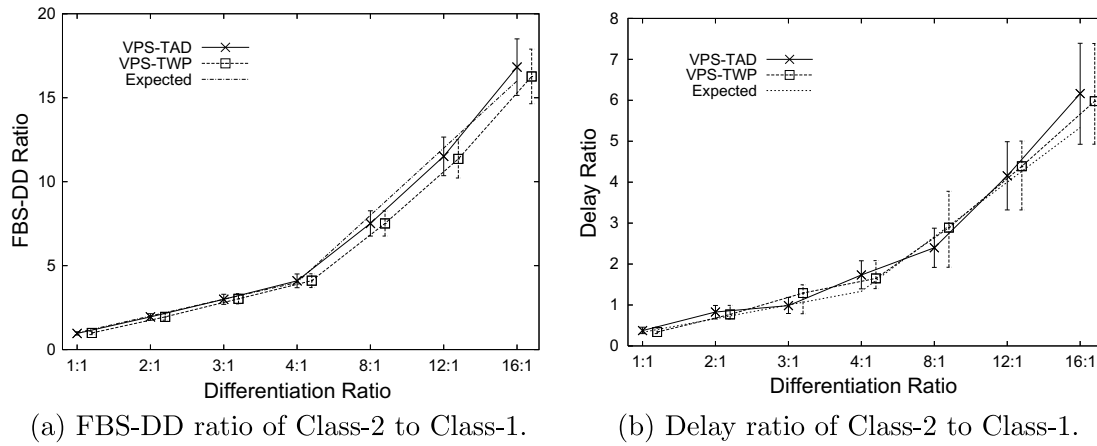


Fig. 7. The fair bandwidth sharing and delay differentiation ratio due to VPS packet scheduling.

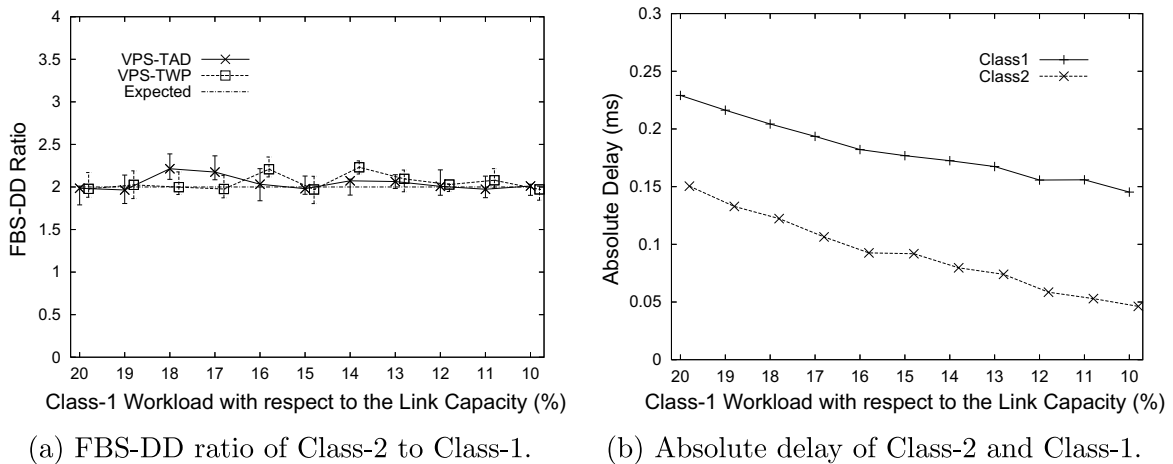


Fig. 8. The fair bandwidth sharing and delay differentiation due to VPS packet scheduling.

proportional loss rate differentiation. The differentiation weight ratio of two classes ($\delta_1 : \delta_2$) is set to 1: 3. The Class-1's workload is fixed at 20% of the link capacity and the Class-2's workload changes from 80% to 160% of the link capacity. Fig. 9(a) shows the achieved FBS-DD ratio by the use of VPS-TAD packet scheduling algorithm with and without the PID control based buffer man-

agement, respectively. The results show that with the control based buffer management, the VPS-TAD scheme is able to achieve more consistent and desirable FBS-DD ratios with respect to both the mean and the variance. Without the control based buffer management, the variance of the FBS-DD ratio is much higher because a class with some bursty traffic can saturate the buffer easily, leaving

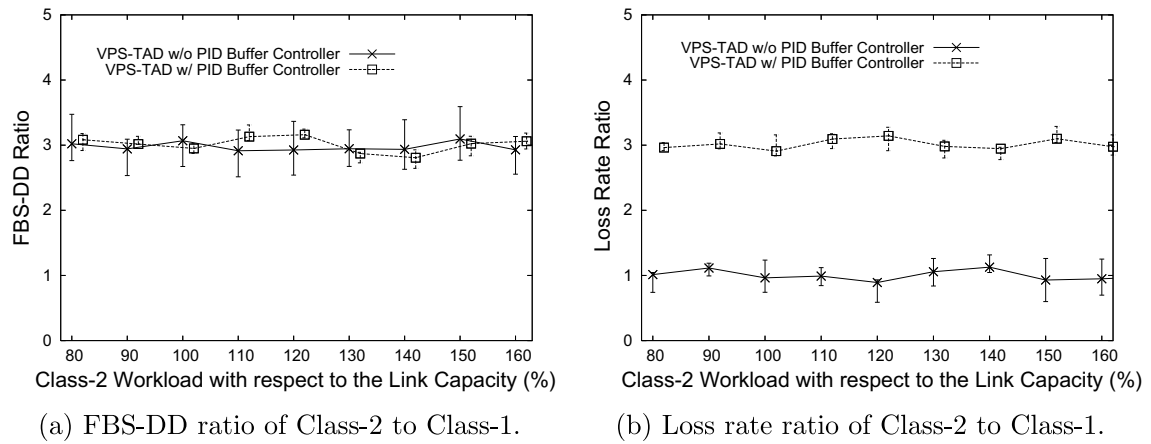


Fig. 9. The impact of the PID control-theoretic buffer management on FBS-DD provisioning.

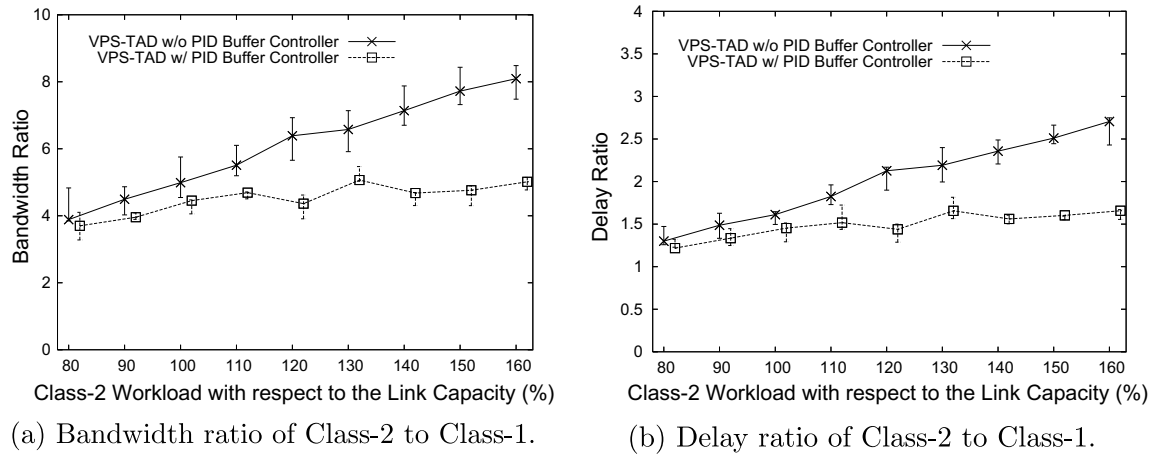


Fig. 10. The impact of the PID control-theoretic buffer management on FBS-DD provisioning.

little buffer space for another class. The VPS-TAD scheduling algorithm aims to minimize the normalized average delays. Its capability is limited by the availability of packets from certain classes for scheduling. This benefits the low-priority but high-workload class. Therefore, the buffer management should be integrated with packet scheduling for controllable FBS-DD provisioning. The integrated approach is capable of self-adapting to varying workloads from different classes, which automatically builds a firewall around aggressive clients and hence protects network resources from saturation. The bandwidth ratio and the delay ratio of Class-2 to Class-1 are further illustrated in Fig. 10.

Fig. 9(b) shows the impact of the PID control-theoretic buffer management on the proportional loss rate differentiation. It shows that with the buffer management, the loss rate ratio of two classes is fairly proportional to the differentiation weight ratio at different workload situations. On the other hand, without the buffer management, both classes experience almost the same loss rate since the packet scheduling algorithm considers the throughput and delay only. This illustrates that the integrated approach is able to achieve the FBS-DD and proportional DiffServ in terms of loss rate simultaneously.

4.3. Three-class fair bandwidth sharing and delay differentiation

We also conducted experiments to study the performance evaluation with three classes. Note again that the number of classes for DiffServ often varies from 2 to 3 [5,7,9,10,14,24,25]. The experi-

mental results by using the VPS-TAD scheme are illustrated in Fig. 11. The differentiation weight ratio of three classes ($\delta_1 : \delta_2 : \delta_3$) is set to 1:2:4 and their workload ratio is 1:1:1. The overall workload changes from 55% to 100%. Results show the approach is able to achieve the FBS-DD ratio and delay ratio as expected. We note that performed a wide range of sensitivity analysis. Due to the space limitation, we only presented those representative results. We note that we did not reach any significantly different conclusions regarding to the predictability and controllability of FBS-DD provisioning achieved by the use of packet scheduling algorithms with the PID control based buffer management.

5. Conclusion

In this paper, we investigated the problem of multi-dimensional QoS differentiation provisioning with respect to both packet delay and bandwidth sharing. We proposed a generalized model, FBS-DD, for providing fair bandwidth sharing with delay differentiation at the same time. One uniqueness is that the delay differentiation and loss rate differentiation are integrated with traffic policing capabilities for providing better controllability to network operators and more fairness to clients. We designed packet scheduling schemes that take both packet delay and packet size into scheduling considerations, without assuming admission control. Furthermore, we designed a PID control-theoretic buffer management scheme. We conducted comprehensive performance evaluation of the schemes with a wide range of sensitivity analysis. Experiment

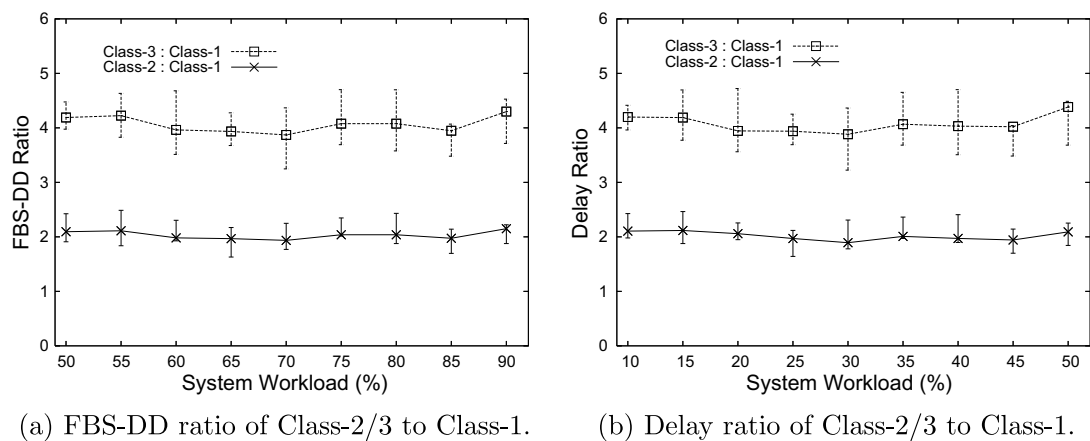


Fig. 11. The impact of the PID control-theoretic buffer management on FBS-DD provisioning.

results by the use of the Bell Labs-I Internet trace files have shown that the FBS-DD model is superior for DiffServ provisioning at dynamic workload conditions. The size-based packet scheduling algorithms can achieve the FBS-DD provisioning at different workload conditions. Results have also demonstrated the significance of the feedback control based buffer management on the performance controllability at overload conditions. The integrated approach can enhance the flexibility of network resource management and enable the multi-dimensional QoS provisioning for popular Internet services.

Differentiated services can be studied in different angles, from the different packet types to the different packet sources. Most of existing delay and loss rate models are studied from the viewpoint of different packet sources, with the rationale of differentiated pricing. Our work has the same viewpoint. It might be interesting to further study the differentiation problem based on different types of packets.

Acknowledgement

A preliminary version of this paper was presented in the Proceedings of the 2007 IEEE International Conference on Computer Communications and Networks. This research was supported in part by U.S. National Science Foundation under Grant No. CNS-0720524 and a CRCW internal grant from the University of Colorado at Colorado Springs. The authors are also grateful to the anonymous reviewers for their valuable comments.

References

- [1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, An architecture for differentiated services, IETF RFC 2475 (1998).
- [2] A. Demers, S. Keshav, and S. Shenker, Analysis and simulation of a fair queueing algorithm, in: Proc. ACM SIGCOMM, 1989, pp. 1–12.
- [3] C. Dovrolis and P. Ramanathan, Proportional differentiated services, part ii: Loss rate differentiation and packet dropping, in: Proc. of the Int'l Workshop on Quality of Service (IWQoS), 2000.
- [4] C. Dovrolis, D. Stiliadis, and P. Ramanathan, Proportional differentiated services: Delay differentiation and packet scheduling, in: Proc. ACM SIGCOMM, 1999, pp. 109–120.
- [5] C. Dovrolis, D. Stiliadis, P. Ramanathan, Proportional differentiated services: Delay differentiation and packet scheduling, IEEE/ACM Trans. on Networking 10 (1) (2002) 12–26.
- [6] L. Essafi, G. Bolch, and A. Andres, An adaptive waiting time priority scheduler for the proportional differentiation model, in: Proc. of the High Performance Computing Symposium, 2001.
- [7] Y. Huang and R. Gu, A simple fifo-based scheme for differentiated loss guarantees, in: Proc. of Int'l Workshop on Quality of Service (IWQoS), 2004.
- [8] E. Kohler, R. Morris, B. Chen, J. Jannotti, M.F. Kaashoek, The click modular router, ACM Trans. on Computer Systems 18 (3) (2000) 263–297.
- [9] Y.-C. Lai, Two schedulers to provide delay proportion and reduce queueing delay simultaneously, Computer Networks: The International Journal of Computer and Telecommunications Networks 51 (11) (2007) 3220–3231.
- [10] L. Le, K. Jeffay, and F.D. Smith, A loss and queueing-delay controller for router buffer management, in: Proc. IEEE Int'l Conf. on Distributed Computing Systems (ICDCS), 2006.
- [11] M.K.H. Leung, J.C.S. Lui, D.K.Y. Yau, Adaptive proportional delay differentiated services: Characterization and performance evaluation, IEEE/ACM Trans. on Networking 9 (6) (2001) 817–908.
- [12] K.C. Leung, V.O.K. Li, Generalized load sharing for packet-switching networks I: Theory and packet-based algorithm, IEEE Trans. on Parallel and Distributed Systems 17 (7) (2006) 694–702.
- [13] C.-C. Li, S.-L. Tsao, M.-C. Chen, Y. Sun, and Y.-M. Huang, Proportional delay differentiation service based on weighted fair queueing, in: Proc. of IEEE Int'l Conf. on Computer Communications and Network (ICCCN), 2000, pp. 418–423.
- [14] J. Liebeherr and N. Christin, JoBS: Joint buffer management and scheduling for differentiated services, in: Proc. Int'l Workshop on Quality of Service (IWQoS), 2001, pp. 404–418.
- [15] Y. Lu, T.F. Abdelzaher, A. Saxena, Design, implementation, and evaluation of differentiated caching services, IEEE Trans. on Parallel and Distributed Systems 15 (5) (2004) 440–452.
- [16] L. Massoulié, J. Roberts, Bandwidth Sharing: Objectives and Algorithms, IEEE/ACM Trans. on Networking 10 (3) (2002) 320–328.
- [17] A. Michalas, M. Louta, P. Fafali, G. Karetsos, V. Loumos, Proportional delay differentiation provision by bandwidth adaptation of class-based queue scheduling, International Journal of Communication Systems 17 (7) (2004) 743–761.
- [18] T. Nandagopal, N. Venkitaraman, R. Sivakumar, and V. Bharghavan, Delay differentiation and adaptation in core stateless networks, in: Proceedings of IEEE INFOCOM, 2000, pp. 421–430.
- [19] J. Nielsen, PMA: Special traces archive. <http://pma.nlanr.net/Special/>.
- [20] A.K. Parekh, R.G. Gallager, A generalized processor sharing approach to flow control in Integrated Services networks: the single-node case, IEEE/ACM Trans. on Networking 1 (3) (1993) 344–357.
- [21] S. Wang, D. Xuan, R. Bettati, W. Zhao, Providing absolute differentiated services for real-time applications in static-priority scheduling networks, IEEE/ACM Trans. on Networking 12 (2) (2004) 326–339.
- [22] J. Wei, C.-Z. Xu, X. Zhou, Q. Li, A robust packet scheduling algorithm for proportional delay differentiation services, Computer Communications, Elsevier 29 (18) (2006) 3679–3690.
- [23] X. Zhou, D. Ippoliti, and T. Boulton, HPPD: A hop-count based probabilistic packet dropper, in: Proc. IEEE Int'l Conf. on Communications (ICC), 2006.
- [24] X. Zhou, J. Wei, C.-Z. Xu, Resource allocation for session-based two-dimensional service differentiation on e-commerce servers, IEEE Trans. on Parallel and Distributed Systems 17 (8) (2006) 838–850.
- [25] X. Zhou, J. Wei, C.-Z. Xu, Quality-of-service differentiation on the internet: A taxonomy, Journal of Network and Computer Applications, Elsevier 30 (1) (2007) 354–383.